LEARNING TARGET-ORIENTED DUAL ATTENTION FOR ROBUST RGB-T TRACKING

Rui Yang, Yabin Zhu, Xiao Wang, Chenglong Li, Jin Tang

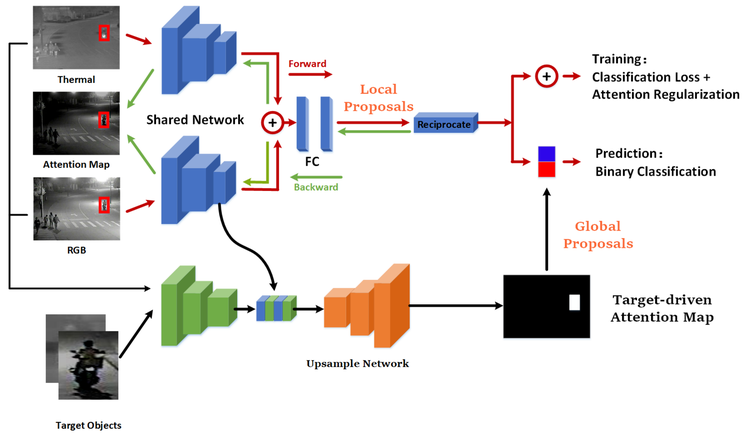
School of Computer Science and Technology, Anhui University, Hefei, China

BACKGROUND & MOTIVATION

- RGB-Thermal object tracking attempts to locate target object using complementary visual and thermal infrared data
- How to learn discriminative and robust features is an import topic of RGBT tracking
- We propose target-oriented dual-attention mechanism for robust RGBT tracking, which includes local attention and global attention.

OUR TRACKER

- The pipeline of our tracker



THE LOCAL ATTENTION

- We backward propagate the score map to obtain two attention maps for each proposal, including positive attention map A_p and negative attention map A_n , then calculating the regularization terms for positive(y=1) and negative(y=0) samples:

$$R_{(y=1)} = \frac{\sigma_{A_p}}{\mu_{A_p}} + \frac{\mu_{A_n}}{\sigma_{A_n}} \qquad R_{(y=0)} = \frac{\mu_{A_p}}{\sigma_{A_p}} + \frac{\sigma_{A_n}}{\mu_{A_n}}$$

- We add the regularization term to the training loss to train the network：

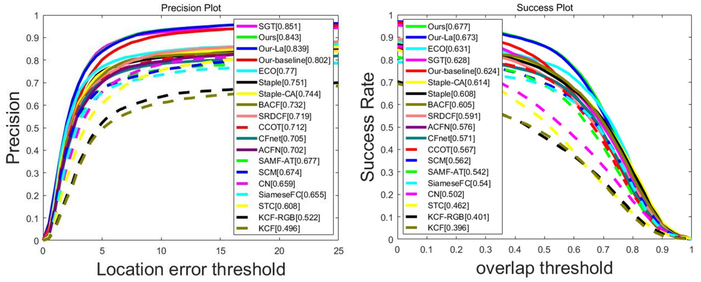
$$\mathrm{L} = L_c + \lambda * [\mathrm{y} * R_{(y=1)} + (1-y) * R_{(y=0)}]$$

where L_c is the cross entropy loss

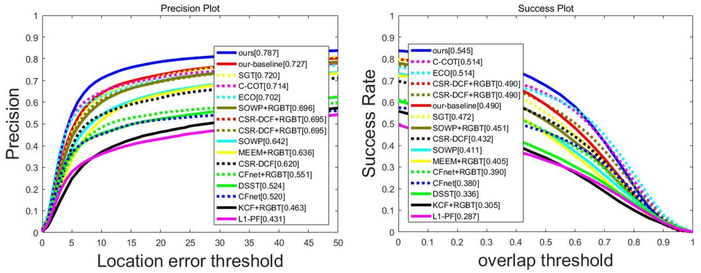
THE GLOBAL ATTENTION

- The limitation of local search strategy: When the target is lost, it's difficult for the tracker to find it back
- To address this limitation, we propose a global proposal generating network to generate the proposals from the whole image

EXPERIMENTS



(a). The tracking results of GTOT-50 dataset.



(b). The tracking results of RGBT-234 dataset.

Table 1. The tracking results with different values of λ.

λ	1	2	3	4	5	6	7	8	9
PR	84.3	84.9	84.3	84.8	83.5	84.1	83.8	84.6	84.7
SR	67.4	67.4	67.7	67.4	66.9	67.5	67.1	66.9	67.7